DOCUMENT RESUME

ED 171 748 TM 008 835

AUTHOR Woodcock, Richard W.
TITLE Woodcock-Johnson Psycho-Educational Battery. Technical Report.
INSTITUTION Teaching Resources Corp., Boston, Mass.
PUB DATE 77
NOTE 14p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Achievement Tests; Adults; Aptitude Tests; *Cognitive Tests; Elementary Secondary Education; Interest Tests; *National Norms; Postsecondary Education; *Predictive Ability (Testing); Preschool Education; *Test Reliability; *Test Validity
IDENTIFIERS *Woodcock Johnson Psychoeducational Battery

ABSTRACT

The Woodcock-Johnson Psycho-Educational Battery is a wide-range comprehensive set of individually administered tests, with norms provided from the preschool to the geriatric level. Part one contains twelve subtests of academic ability and aptitude. Part two contains ten subtests measuring academic achievement. The final section measures scholastic and nonscholastic interests. Norms were based upon a nationwide sample of 4,732 subjects, proportioned to match the exact distribution of sex-by-race, region-by-urbanization, and occupational status of the U.S. population. Concurrent validity was gathered by comparing results of standardized tests administered concurrently with the reading, mathematics, written language, and knowledge aptitude subtests--from part one of the battery. These same subtests were also used to predict achievement on the corresponding achievement subtests in the battery. (Fifteen tables report norms, reliability, and validity--mainly as cluster scores). (CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED171748

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Woodcock-Johnson Psycho-Educational Battery

Technical Report

Richard W. Woodcock

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Teaching Resources Corporation

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Teaching Resources
A New York Times Company

TM008 835

The purpose of this Technical Report is to provide potential users and others with the major information regarding norming, reliability, and validity of the battery. A technical manual is in preparation which contains detailed information on the research and development underlying the Woodcock-Johnson Psycho-Educational Battery.

© 1977 by Teaching Resources Corporation
100 Boylston Street
Boston, Massachusetts 02116

DESCRIPTION

The *Woodcock-Johnson Psycho-Educational Battery* is a wide-range comprehensive set of tests for measuring cognitive ability, achievement, and interest. The tests are individually administered and norms are provided from the preschool to the geriatric level. Part One of the battery contains 12 subtests. Scores from clusters of these subtests provide information regarding a subject's cognitive ability and scholastic aptitudes. Part Two contains 10 subtests measuring several aspects of scholastic achievement. Part Three contains five subtests measuring level of preference for participating in various scholastic and non-scholastic forms of activity.

The primary unit of interpretation for the *Woodcock-Johnson Psycho-Educational Battery* is the cluster score. Each cluster is made up of two or more subtests. Table 1 shows the composition of all clusters. The table also indicates whether each cluster is a weighted or unweighted (weights of 1.0) composite. Weighted composites are based on the results of multiple regression and principal component analysis studies.

ITEM ANALYSIS AND CALIBRATION

The development of the battery followed traditional stages of determining specifications, developing subtest procedures and test items, and proceeding through a series of small scale item analysis and editing stages. The concepts of latent trait theory and the analysis of data by the Rasch model were extensively employed (Hambleton & Cook, 1977; Rasch, 1960; Wright, 1968; and Wright, 1977). The "relative performance index" and "instructional range" used in the interpretation of battery scores are special transformations and extensions of the Rasch log ability scale (Woodcock, 1973; Woodcock & Dahl, 1971).

NORMING

Normative data were gathered from 4732 subjects in 42 communities widely distributed throughout the conterminous United States. All data in the school-age sample were gathered throughout a one school year period extending from April, 1976 to March, 1977. Adult data were gathered from April, 1976 until May, 1977.

TABLE 1. SUBTEST COMPOSITION OF CLUSTERS

CLUSTER	COMBINATION OF SUBTESTS[1]	
	Weighted	Unweighted
Part One: Tests of Cognitive Ability		
Broad Cognitive Ability		
Full Scale	1 to 12	–
Preschool Scale	1 to 6	–
Brief Scale	6, 8	–
Scholastic Aptitudes		
Reading Aptitude	4, 5, 8, 12	–
Mathematics Aptitude	7, 8, 9, 11	–
Written Language Aptitude	6, 7, 8, 10	–
Knowledge Aptitude	3, 6, 8, 12	–
Part Two: Tests of Achievement		
Reading	–	13, 14, 15
Mathematics	–	16, 17
Written Language	–	18, 19
Knowledge	–	20, 21, 22
Skills (Preschool)	–	13, 17, 18
Part Three: Tests of Interest Level		
Scholastic Interest	–	23, 24, 25
Nonscholastic Interest	–	26, 27

[1]See Table 4 for subtest names.

Table 2 reports the distribution of the norming sample by selected levels in respect to the characteristics of sex-by-race and region-by-urbanization. Table 3 reports the distribution of the norming sample in respect to occupational status for subjects 20 years of age or older.

Each subject in the norming sample was assigned a weight based upon that person's required contribution to a data base proportioned to match the exact distributions of sex-by-race, region-by-urbanization, and occupational status in the total U.S. population. If a subject belonged to any subcategory (such as "white female" or "urban South") that had a larger proportion in the norming sample than the proportion in the U.S. population at that age level, then this subject was assigned some weight less than "1." On the other hand, the subject would receive some weight greater than "1" if belonging to a subcategory having a lower proportion in the sample than in the U.S. population. Estimates of the distribution of the U.S. population in respect to sex-by-race were based upon July 1, 1976 estimates (U.S. Bureau of the Census, 1977). The latest available information for the region-by-urbanization distribution is provided in the 1970 census reports (U.S. Bureau of the Census, 1971). Subjects who were 20 years of age or older received a third weight based upon the extent to which their occupational classification may have been under- or overrepresented in respect to the U.S. population in 1976 (U.S. Department of Labor, 1977). The final norming weight assigned each subject was a product of the two or three weights obtained in the manner described above.

TABLE 2. DISTRIBUTION OF NORMING SAMPLE AND U.S. POPULATION BY LEVEL FOR SEX-BY-RACE AND FOR REGION-BY-URBANIZATION

LEVEL	NUMBER IN SAMPLE	PERCENT											
		Sex-by-Race				Region-by-Urbanization							
		White Male	White Female	Nonwhite Male	Nonwhite Female	Urban Northeast	Nonurban Northeast	Urban North Central	Nonurban North Central	Urban South	Nonurban South	Urban West	Nonurban West
Age 3 to 5													
Sample:	555	39.3	43.8	8.3	8.6	10.0	1.8	15.7	14.4	24.7	2.7	21.8	8.8
US:		42.4	40.2	8.9	8.7	14.4	5.9	14.9	11.9	15.6	18.6	12.6	6.0
Grade K to 12													
Sample:	3935	41.1	46.6	5.6	6.6	13.1	3.4	15.1	22.8	12.4	6.2	15.9	11.2
US:		42.9	41.0	8.1	8.0	16.0	6.5	15.3	12.2	14.6	17.6	12.1	5.7
Age 18 to 64													
Sample:	503	40.2	43.9	6.0	9.9	19.1	2.0	8.0	33.0	19.3	3.6	9.1	6.0
US:		43.2	44.5	6.6	5.7	16.6	6.8	14.8	11.8	14.5	17.3	12.4	5.9
Age 65+													
Sample:	97	37.1	55.7	2.1	5.2	16.5	1.0	7.2	24.7	25.8	1.0	11.3	12.4
US:		36.9	53.9	4.0	5.2	17.5	7.2	15.1	12.0	14.6	17.4	11.0	5.2

TABLE 3. DISTRIBUTION OF NORMING SAMPLE AND U.S. POPULATION BY AGE FOR OCCUPATIONAL STATUS

LEVEL	NUMBER IN SAMPLE	PERCENT					
		Not in Labor Force	Unemployed	White Collar	Blue Collar	Farm Workers	Service Workers
Age 20 to 64							
Sample:	351	29.1	3.1	35.9	21.1	1.1	9.7
US:		27.9	4.6	34.8	22.3	2.1	8.3
Age 65+							
Sample:	97	93.8	0.0	3.1	1.0	0.0	2.1
US:		86.8	0.7	6.5	4.1	0.4	1.5

RELIABILITY

Table 4 reports the reliability of subtest scores across their range of use. The numbers in parentheses below each reliability coefficient indicate the number of subjects involved in that calculation. With the exception of five subtests, which have been identified in the table, all reliabilities are based upon split-half correlations of the odd-even raw scores, and corrected by the Spearman-Brown formula. Subtests 2 and 7 are timed tests and the reliability coefficients reported are test-retest correlations. All retest scores were obtained at least one but not more than two days after the first test was administered. Subtests 4, 9, and 11, are tests which may be discontinued at specified points if the subject is doing poorly. The nature of these three tests, however, is such that a subject would be expected to obtain a few more points if the test had been continued. Separate log ability scoring tables, with associated error values, were obtained through Rasch calibration procedures for each length of test where a discontinuance was allowed. The standard error of measurement associated with each subject's score, and for the length of test administered, was then determined and the reliability of the subtest scores calculated by

$$r = 1 - \frac{\Sigma SE_m^2 \div N}{SD^2}$$

where SE_m is the error term associated with a given score for that length of test, N is the number of subjects in the reliability sample, and SD is the standard deviation of log ability scores for the reliability sample.

The reliability of both weighted and unweighted cluster scores was calculated by the following formula (Guilford, 1954):

$$r = 1 - \frac{\Sigma w^2{}_j \sigma^2{}_j - \Sigma w^2{}_j \sigma^2{}_j r_{jj}}{\Sigma w^2{}_j \sigma^2{}_j + 2\Sigma w_j w_k \sigma_j \sigma_k r_{jk}}$$

where w is the weight assigned to any subtest, σ is the standard deviation of any subtest, r_{jj} is the reliability of any subtest, and r_{jk} is the correlation between two subtests. Table 5 reports the cluster score reliabilities.

VALIDITY

Concurrent validity data have been gathered from four samples of normal subjects. A sample of 40 four-year-olds attending a day care center was administered the following tests:

Denver Developmental Screening Test (Frankenburg, 1969)
Developmental Indicators for the Assessment of Learning (DIAL) (Mardell & Goldenberg, 1975)
Peabody Picture Vocabulary Test (Dunn, 1965)
Stanford-Binet Intelligence Scale, Form L-M (Terman & Merrill, 1973)
Woodcock-Johnson Psycho-Educational Battery (preschool portions)

A group of 83 third-grade students and a group of 86 fifth-grade students were administered the following tests:

Iowa Tests of Basic Skills, Form 5 (Hieronymous & Lindquist, 1971)
KeyMath Diagnostic Arithmetic Test (Connolly, Natchman, & Pritchett, 1976)
Peabody Individual Achievement Test (Dunn & Markwardt, 1970)
Peabody Picture Vocabulary Test
Wechsler Intelligence Scale for Children - Revised (WISC-R) (Wechsler, 1974)
Wide Range Achievement Test (Jastak, Bijou & Jastak, 1965)
Woodcock Reading Mastery Tests (Woodcock, 1973)
Woodcock-Johnson Psycho-Educational Battery

TABLE 4. SUBTEST SCORE RELIABILITIES

SUBTEST	LEVEL										
	Age 3	Age 4	K	Grade 1	Grade 3	Grade 5	Grade 8	Grade 12	Age 20-39	Age 40-64	Age 65+
Part One:											
1. Picture Vocabulary	.74 (144)	.69 (147)	.81 (279)	.82 (522)	.79 (471)	.79 (597)	.85 (475)	.86 (398)	.89 (162)	.92 (185)	.93 (91)
2. Spatial Relations[1]	.86 (32)	.57 (21)	.90 (41)	.88 (79)	.89 (56)	.81 (61)	.79 (87)	.87 (54)	–	–	–
3. Memory for Sentences	.86 (146)	.84 (147)	.85 (277)	.83 (521)	.80 (472)	.77 (596)	.76 (475)	.72 (401)	.76 (161)	.77 (182)	.90 (86)
4. Visual-Auditory Learning[2]	.95 (143)	.95 (149)	.96 (276)	.95 (520)	.87 (471)	.85 (595)	.87 (470)	.88 (397)	.96 (162)	.99 (180)	.91 (87)
5. Blending	.91 (144)	.87 (146)	.88 (276)	.89 (522)	.86 (471)	.85 (596)	.87 (475)	.90 (401)	.92 (160)	.95 (182)	.93 (86)
6. Quantitative Concepts	.72 (145)	.80 (146)	.81 (279)	.83 (522)	.85 (471)	.84 (596)	.89 (475)	.90 (402)	.92 (161)	.92 (184)	.93 (90)
7. Visual Matching[1]	–	–	–	.84 (78)	.79 (56)	.61 (50)	.63 (88)	.65 (56)	–	–	–
8. Antonyms-Synonyms	–	–	–	.89 (523)	.87 (471)	.86 (597)	.90 (474)	.88 (401)	.91 (161)	.93 (185)	.94 (91)
9. Analysis-Synthesis[2]	–	–	–	.86 (519)	.93 (471)	.78 (598)	.86 (471)	.84 (396)	.85 (161)	.86 (184)	.82 (89)
10. Numbers Reversed	–	–	–	.73 (522)	.70 (472)	.72 (596)	.81 (474)	.83 (399)	.89 (161)	.88 (182)	.86 (87)
11. Concept Formation[2]	–	–	–	.89 (523)	.92 (471)	.91 (596)	.87 (473)	.83 (396)	.90 (161)	.94 (184)	.91 (87)
12. Analogies	–	–	–	.80 (523)	.80 (471)	.84 (596)	.84 (475)	.86 (400)	.85 (160)	.89 (184)	.94 (91)
Part Two:											
13. Letter-Word Identification	.89 (145)	.89 (149)	.97 (280)	.97 (521)	.92 (472)	.90 (597)	.95 (475)	.93 (416)	.97 (162)	.97 (186)	.96 (89)
14. Word Attack	–	–	–	.94 (521)	.92 (472)	.92 (597)	.91 (474)	.89 (415)	.91 (163)	.91 (185)	.86 (90)
15. Passage Comprehension	–	–	–	.92 (521)	.86 (471)	.80 (594)	.82 (474)	.73 (414)	.86 (161)	.89 (185)	.87 (88)
16. Calculation	–	–	–	.90 (523)	.88 (472)	.87 (597)	.84 (473)	.88 (414)	.91 (162)	.94 (185)	.94 (88)
17. Applied Problems	.85 (144)	.67 (148)	.89 (280)	.85 (521)	.87 (471)	.84 (597)	.83 (473)	.87 (412)	.90 (162)	.90 (185)	.88 (86)
18. Dictation	.46 (144)	.70 (147)	.81 (279)	.87 (522)	.90 (472)	.88 (596)	.90 (474)	.88 (415)	.92 (162)	.93 (186)	.96 (89)
19. Proofing	–	–	–	.89 (519)	.86 (468)	.90 (592)	.90 (472)	.88 (413)	.91 (161)	.95 (185)	.96 (88)
– Punctuation & Capitalization	–	–	–	.60 (519)	.76 (468)	.77 (590)	.83 (472)	.77 (412)	.80 (155)	.88 (155)	.94 (88)
– Spelling	–	–	–	.86 (519)	.88 (468)	.88 (590)	.90 (472)	.87 (412)	.92 (155)	.94 (155)	.95 (88)
– Usage	–	–	–	.87 (519)	.81 (468)	.79 (590)	.81 (472)	.76 (412)	.80 (155)	.90 (155)	.93 (88)
20. Science	.82 (145)	.85 (148)	.85 (279)	83 (521)	.78 (471)	.82 (596)	.84 (473)	.86 (415)	.86 (161)	.89 (184)	.88 (92)
21. Social Studies	.73 (145)	.83 (149)	.81 (279)	.80 (521)	.81 (471)	.78 (595)	.87 (473)	.84 (415)	.87 (162)	.86 (184)	.87 (90)
22. Humanities	.84 (145)	.77 (147)	.84 (279)	.78 (522)	.77 (471)	.81 (595)	.90 (473)	.86 (415)	.84 (162)	.93 (184)	.92 (89)
Part Three:											
23. Reading Interest	–	–	–	–	–	.81 (590)	.86 (474)	.88 (416)	.90 (161)	.88 (153)	.85 (89)
24. Mathematics Interest	–	–	–	–	–	.87 (590)	.86 (474)	.86 (416)	.88 (161)	.89 (183)	.86 (89)
25. Written Language Interest	–	–	–	–	–	.84 (590)	.83 (474)	.88 (416)	.87 (161)	.89 (183)	.92 (89)
26. Physical Interest	–	–	–	–	–	.75 (590)	.79 (474)	.85 (417)	.91 (160)	.89 (185)	.91 (90)
27. Social Interest	–	–	–	–	–	.73 (590)	.78 (474)	.80 (417)	.86 (160)	.84 (185)	.78 (90)

[1]Timed tests. Test-retest reliabilities reported.

[2]Tests with cut-offs. Generic reliabilities reported.

TABLE 5. CLUSTER SCORE RELIABILITIES

CLUSTER	LEVEL										
	Age 3	Age 4	K	Grade 1	Grade 3	Grade 5	Grade 8	Grade 12	Age 20-39	Age 40-64	Age 65+
Part One:											
Broad Cognitive Ability											
Full Scale	–	–	–	.97 (512)	.97 (468)	.96 (592)	.97 (472)	.97 (389)	.98 (155)	.98 (180)	.98 (83)
Brief Scale	–	–	–	.91 (522)	.92 (471)	.91 (596)	.94 (474)	.93 (397)	.95 (160)	.96 (184)	.98 (90)
Preschool Scale	.93 (140)	.92 (142)	.94 (274)	.94 (517)	–	–	–	–	–	–	–
Scholastic Aptitudes											
Reading Aptitude	–	–	–	.94 (518)	.93 (470)	.93 (594)	.95 (474)	.95 (391)	.96 (156)	.97 (180)	.97 (85)
Mathematics Aptitude	–	–	–	.94 (517)	.92 (469)	.86 (596)	.88 (473)	.86 (392)	.87 (157)	.90 (181)	.91 (86)
Written Language Aptitude	–	–	–	.92 (520)	.92 (471)	.90 (595)	.93 (473)	.93 (394)	.96 (159)	.96 (181)	.97 (86)
Knowledge Aptitude	–	–	–	.93 (518)	.94 (471)	.93 (594)	.95 (474)	.95 (393)	.96 (157)	.97 (181)	.98 (85)
Part Two:											
Reading	–	–	–	.98 (521)	.96 (471)	.95 (594)	.96 (474)	.94 (414)	.97 (161)	.97 (185)	.96 (86)
Mathematics	–	–	–	.92 (521)	.92 (471)	.91 (597)	.90 (473)	.93 (412)	.94 (162)	.96 (185)	.95 (85)
Written Language	–	–	–	.92 (519)	.93 (468)	.94 (590)	.94 (472)	.93 (412)	.95 (161)	.97 (185)	.98 (88)
Knowledge	.90 (145)	.92 (145)	.93 (279)	.92 (521)	.91 (471)	.92 (594)	.95 (473)	.94 (415)	.94 (161)	.96 (184)	.96 (89)
Skills (Preschool)	.84 (141)	.86 (145)	.93 (278)	.97 (521)	–	–	–	–	–	–	–
Part Three:											
Scholastic Interest	–	–	–	–	–	.92 (590)	.91 (474)	.93 (416)	.93 (160)	.96 (179)	.96 (89)
Nonscholastic Interest	–	–	–	–	–	.81 (590)	.84 (474)	.87 (417)	.92 (159)	.89 (183)	.90 (90)

A group of 75 twelfth-grade students was administered the following tests:

Peabody Individual Achievement Test
Peabody Picture Vocabulary Test
Wechsler Adult Intelligence Scale (Wechsler, 1955)
Wide Range Achievement Test
Woodcock Reading Mastery Tests
Woodcock-Johnson Psycho-Educational Battery

An additional group of 40 twelfth-grade students was administered the following tests:

Iowa Tests of Educational Development (Lindquist & Feldt, 1972)
Wide Range Achievement Test
Woodcock Reading Mastery Tests
Woodcock-Johnson Psycho-Educational Battery (Parts Two and Three)

Table 6 summarizes the correlations obtained among the measures of cognitive ability administered to the subjects in the four validity samples. Tables 7 through 10 report concurrent validity coefficients for tests of reading, mathematics, written language, and knowledge respectively.

TABLE 6. CONCURRENT VALIDITY OF COGNITIVE ABILITY TESTS

	Psycho-Ed Cognitive Ability			Denver	DIAL	PPVT	Stanford-Binet	Wechsler		
	Full Scale	Preschool Scale	Brief Scale					Full Scale	Verbal	Performance
Preschool Study: (N=40)										
Psycho-Ed Preschool	–	1.00	–	.54	.60	.74	.83	–	–	–
Denver	–	.54	–	1.00	.82	.52	.65	–	–	–
DIAL	–	.60	–	.82	1.00	.60	.74	–	–	–
PPVT	–	.74	–	.52	.60	1.00	.73	–	–	–
Stanford-Binet	–	.83	–	.65	.74	.73	1.00	–	–	–
Grade 3 Study: (N=83)										
Psycho-Ed Full Scale	1.00	–	.91	–	–	.58	–	.79	.76	.62
Psycho-Ed Brief Scale	.91	–	1.00	–	–	.56	–	.77	.81	.51
PPVT	.58	–	.56	–	–	1.00	–	.60	.60	.37
WISC-R FS	.79	–	.77	–	–	.60	–	1.00	.92	.84
WISC-R V	.76	–	.81	–	–	.60	–	.92	1.00	.55
WISC-R P	.62	–	.51	–	–	.37	–	.84	.55	1.00
Grade 5 Study: (N=86)										
Psycho-Ed Full Scale	1.00	–	.84	–	–	.64	–	.79	.75	.58
Psycho-Ed Brief Scale	.84	–	1.00	–	–	.67	–	.67	.72	.39
PPVT	.64	–	.67	–	–	1.00	–	.52	.64	.22
WISC-R FS	.79	–	.67	–	–	.52	–	1.00	.86	.82
WISC-R V	.75	–	.72	–	–	.64	–	.86	1.00	.42
WISC-R P	.58	–	.39	–	–	.22	–	.82	.42	1.00
Grade 12 Study: (N=75)										
Psycho-Ed Full Scale	1.00	–	.87	–	–	.56	–	.83	.81	.67
Psycho-Ed Brief Scale	.87	–	1.00	–	–	.62	–	.82	.85	.60
PPVT	.56	–	.62	–	–	1.00	–	.59	.63	.40
WAIS FS	.83	–	.82	–	–	.59	–	1.00	.93	.89
WAIS V	.81	–	.85	–	–	.63	–	.93	1.00	.66
WAIS P	.67	–	.60	–	–	.40	–	.89	.66	1.00

TABLE 7. CONCURRENT VALIDITY OF READING TESTS

	Psycho-Ed	Iowa	PIAT (2+3)	WRAT	WRMT
Grade 3 Study: (N=83)					
Psycho-Ed	1.00	.81	.91	.89	.92
Iowa	.81	1.00	.78	.82	.83
PIAT(2+3)	.91	.78	1.00	.88	.87
WRAT	.89	.82	.88	1.00	.88
WRMT	.92	.83	.87	.88	1.00
Grade 5 Study: (N=86)					
Psycho-Ed	1.00	.76	.75	.84	.87
Iowa	.76	1.00	.65	.76	.78
PIAT(2+3)	.75	.65	1.00	.75	.78
WRAT	.84	.76	.75	1.00	.86
WRMT	.87	.78	.78	.86	1.00
Grade 12 Study: (N=75)					
Psycho-Ed	1.00	.78	.86	.90	.90
Iowa (N=40)	.78	1.00	–	.75	.79
PIAT(2+3)	.86	–	1.00	.87	.87
WRAT	.90	.75	.87	1.00	.90
WRMT	.90	.79	.87	.90	1.00

The prediction of achievement from aptitude. A major function of the *Woodcock-Johnson Psycho-Educational Battery* is to provide statements regarding a person's expected achievement level in four areas of achievement: reading, mathematics, written language, and knowledge. Special aptitude clusters have been provided for making such predictions. The four subtests included in each aptitude cluster were selected primarily on the basis of multiple regression studies. A subtest was not considered as a potential predictor if it was believed that the subtest had common content with the criterion. For example, Quantitative Concepts was not included in the mathematics aptitude cluster even though its correlation with mathematics achievement was quite high. Picture Vocabulary was not included in the knowledge aptitude cluster although that subtest correlated highly with the knowledge criterion. Furthermore, though it correlated well with reading achievement, the Quantitative Concepts subtest was not included in the reading aptitude cluster because it was believed some subjects might obtain low scores on the Quantitative Concepts subtest primarily because they have difficulty in reading and perhaps have been retained in grade one or more times. If such were the case, and Quantitative Concepts were included in the reading aptitude cluster, then a subject might receive a lower reading aptitude score primarily because of lack of opportunity to learn quantitative concepts.

Tables 11 through 15 summarize the correlations between the aptitude cluster scores and their respective achievement criteria. In addition, these tables include comparative results with other measures that are sometimes used to determine expected achievement. If a coefficient in one of these tables may be spuriously high due to the presence of common content in both the predictor and criterion measures, this has been indicated by an asterisk.

The prediction of end of first grade achievement. Cognitive Ability scores, obtained at the end of kindergarten (N=42), correlated .67 with Reading Achievement scores and .75 with Knowledge scores obtained 12 months later at the end of first grade. Cognitive Ability scores, obtained at the beginning of grade one (N=73), correlated .67 with Reading Achievement scores and .75 with Knowledge scores obtained 8 months later at the end of first grade.

TABLE 8. CONCURRENT VALIDITY OF MATHEMATICS TESTS

	Psycho-Ed	Iowa	KeyMath	PIAT	WRAT
Grade 3 Study: (N=83)					
Psycho-Ed	1.00	.62	.82	.70	.46
Iowa	.62	1.00	.75	.68	.41
KeyMath	.82	.75	1.00	.76	.42
PIAT	.70	.68	.76	1.00	.40
WRAT	.46	.41	.42	.40	1.00
Grade 5 Study: (N=86)					
Psycho-Ed	1.00	.77	.80	.69	.78
Iowa	.77	1.00	.78	.68	.74
KeyMath	.80	.78	1.00	.70	.76
PIAT	.69	.68	.70	1.00	.59
WRAT	.78	.74	.76	.59	1.00
Grade 12 Study: (N=75)					
Psycho-Ed	1.00	.71	–	.74	.81
Iowa (N=40)	.71	1.00	–	–	.66
PIAT	.74	–	–	1.00	.68
WRAT	.81	.66	–	.68	1.00

TABLE 9. CONCURRENT VALIDITY OF WRITTEN LANGUAGE TESTS

	Psycho-Ed	Iowa	PIAT	WRAT
Grade 3 Study: (N=83)				
Psycho-Ed	1.00	.84	.78	.83
Iowa	.84	1.00	.72	.75
PIAT	.78	.72	1.00	.75
WRAT	.83	.75	.75	1.00
Grade 5 Study: (N=86)				
Psycho-Ed	1.00	.80	.76	.71
Iowa	.80	1.00	.64	.56
PIAT	.76	.64	1.00	.54
WRAT	.71	.56	.54	1.00
Grade 12 Study: (N=75)				
Psycho-Ed	1.00	.74	.79	.86
Iowa (N=40)	.74	1.00	–	.72
PIAT	.79	–	1.00	.74
WRAT	.86	.72	.74	1.00

TABLE 10. CORRELATION OF PSYCHO-ED KNOWLEDGE WITH PIAT GENERAL INFORMATION

Grade 3 Study: (N=83)	.72
Grade 5 Study: (N=86)	.68
Grade 12 Study: (N=74)	.83

TABLE 11. PREDICTION OF PRESCHOOL ACHIEVEMENT SCORES (N=40)

PREDICTORS	ACHIEVEMENT MEASURES	
	Psycho-Ed Knowledge	Psycho-Ed Skills
Psycho-Ed Preschool Scale	.77*	.45
Denver	.54*	.39
DIAL	.64*	.42
PPVT	.91**	.36
Stanford-Binet	.80*	.47

*Correlation may be inflated due to some common content in the predictor.

**Correlation is probably inflated due to considerable common content in the predictor.

TABLE 12. PREDICTION OF READING ACHIEVEMENT SCORES

PREDICTORS	READING ACHIEVEMENT MEASURES				
	Psycho-Ed	Iowa	PIAT (2+3)	WRAT	WRMT
Grade 3 Study: (N=83)					
Psycho-Ed Reading Aptitude	.78	.76	.73	.77	.85
Psycho-Ed Full Scale	.80	.78	.74	.78	.83
Psycho-Ed Brief Scale	.77	.78	.76	.78	.83
PPVT	.46	.59	.47	.42	.43
WISC-R Full Scale	.67	.72	.61	.66	.70
WISC-R Verbal Scale	.68	.74	.63	.67	.71
WISC-R Performance Scale	.48	.50	.42	.46	.49
Grade 5 Study: (N=86)					
Psycho-Ed Reading Aptitude	.69	.69	.69	.66	.75
Psycho-Ed Full Scale	.71	.71	.69	.67	.75
Psycho-Ed Brief Scale	.67	.69	.69	.64	.71
PPVT	.56	.68	.56	.58	.62
WISC-R Full Scale	.54	.56	.58	.49	.61
WISC-R Verbal Scale	.57	.60	.56	.51	.62
WISC-R Performance Scale	.33	.32	.41	.30	.41
Grade 12 Study: (N=75)					
Psycho-Ed Reading Aptitude	.78	–	.79	.84	.85
Psycho-Ed Full Scale	.78	–	.78	.82	.88
Psycho-Ed Brief Scale	.72	–	.70	.78	.81
PPVT	.58	–	.61	.67	.63
WAIS Full Scale	.66	–	.68	.74	.82
WAIS Verbal Scale	.67	–	.69	.77	.83
WAIS Performance Scale	.49	–	.50	.56	.65

TABLE 13. PREDICTION OF MATHEMATICS ACHIEVEMENT SCORES

PREDICTORS	MATHEMATICS ACHIEVEMENT MEASURES				
	Psycho-Ed	Iowa	KeyMath	PIAT	WRAT
Grade 3 Study: (N=83)					
Psycho-Ed Mathematics Aptitude	.71	.67	.76	.64	.46
Psycho-Ed Full Scale	.75*	.76*	.87*	.69*	.48
Psycho-Ed Brief Scale	.77**	.73**	.85**	.70**	.45
PPVT	.30	.41	.48	.46	.18
WISC-R Full Scale	.59*	.68*	.72*	.63*	.41*
WISC-R Verbal Scale	.62*	.65*	.76*	.62*	.42*
WISC-R Performance Scale	.39	.54	.48	.46	.27
Grade 5 Study: (N=86)					
Psycho-Ed Mathematics Aptitude	.68	.66	.70	.66	.66
Psycho-Ed Full Scale	.77*	.67*	.82*	.68*	.70
Psycho-Ed Brief Scale	.64**	.63**	.72**	.67**	.63
PPVT	.48	.42	.51	.46	.44
WISC-R Full Scale	.61*	.48*	.66*	.54*	.55*
WISC-R Verbal Scale	.62*	.49*	.63*	.56*	.51*
WISC-R Performance Scale	.47	.30	.48	.33	.41
Grade 12 Study: (N=75)					
Psycho-Ed Mathematics Aptitude	.70	–	–	.66	.69
Psycho-Ed Full Scale	.74*	–	–	.79*	.71
Psycho-Ed Brief Scale	.77**	–	–	.77**	.68
PPVT	.53	–	–	.55	.41
WAIS Full Scale	.75*	–	–	.71*	.74*
WAIS Verbal Scale	.71*	–	–	.71*	.68*
WAIS Performance Scale	.62	–	–	.53	.67

*Correlation may be inflated due to some common content in the predictor.
**Correlation is probably inflated due to considerable common content in the predictor.

TABLE 14. PREDICTION OF WRITTEN LANGUAGE ACHIEVEMENT SCORES

PREDICTORS	WRITTEN LANGUAGE ACHIEVEMENT MEASURES			
	Psycho-Ed	Iowa	PIAT	WRAT
Grade 3 Study: (N=83)				
Psycho-Ed Written Language Aptitude	.81	.73	.75	.71
Psycho-Ed Full Scale	.82	.76	.69	.68
Psycho-Ed Brief Scale	.78	.70	.73	.68
PPVT	.41	.36	.41	.32
WISC-R Full Scale	.68	.66	.63	.64
WISC-R Verbal Scale	.67	.61	.66	.67
WISC-R Performance Scale	.51	.56	.41	.42
Grade 5 Study: (N=86)				
Psycho-Ed Written Language Aptitude	.65	.67	.54	.44
Psycho-Ed Full Scale	.67	.69	.57	.49
Psycho-Ed Brief Scale	.62	.63	.50	.45
PPVT	.45	.48	.35	.34
WISC-R Full Scale	.68	.54	.43	.39
WISC-R Verbal Scale	.65	.51	.37	.40
WISC-R Performance Scale	.49	.41	.36	.26
Grade 12 Study: (N=75)				
Psycho-Ed Written Language Aptitude	.71	–	.79	.86
Psycho-Ed Full Scale	.74	–	.59	.71
Psycho-Ed Brief Scale	.63	–	.55	.57
PPVT	.62	–	.52	.50
WAIS Full Scale	.65	–	.54	.63
WAIS Verbal Scale	.64	–	.54	.63
WAIS Performance Scale	.51	–	.41	.51

TABLE 15. PREDICTION OF KNOWLEDGE MEASURES

PREDICTORS	KNOWLEDGE MEASURES	
	Psycho-Ed	PIAT
Grade 3 Study: (N=83)		
Psycho-Ed Knowledge Aptitude	.76	.64
Psycho-Ed Full Scale	.72*	.65
Psycho-Ed Brief Scale	.76	.70
PPVT	.71*	.58
WISC-R Full Scale	.73*	.63*
WISC-R Verbal Scale	.76*	.71*
WISC-R Performance Scale	.49	.36
Grade 5 Study: (N=86)		
Psycho-Ed Knowledge Aptitude	.78	.60
Psycho-Ed Full Scale	.72*	.50
Psycho-Ed Brief Scale	.77	.58
PPVT	.69	.68
WISC-R Full Scale	.64*	.44*
WISC-R Verbal Scale	.75*	.55*
WISC-R Performance Scale	.30	.18
Grade 12 Study: (N=75)		
Psycho-Ed Knowledge Aptitude	.87	.80
Psycho-Ed Full Scale	.84*	.67
Psycho-Ed Brief Scale	.85	.78
PPVT	.64	.66
WAIS Full Scale	.79*	.71*
WAIS Verbal Scale	.82*	.74*
WAIS Performance Scale	.57	.41

*Correlation may be inflated due to some common content in the predictor.

REFERENCES

Connolly, A.J., Nachtman, W., and Pritchett, E.M., *KeyMath Diagnostic Arithmetic Test,* Circle Pines, Minnesota, American Guidance Service, 1971, 1976.

Dunn, L.M., *Peabody Picture Vocabulary Test,* Circle Pines, Minnesota, American Guidance Service, 1965.

Dunn, L.M., and Markwardt, F.C., *Peabody Individual Achievement Test,* Circle Pines, Minnesota, American Guidance Service, 1970.

Frankenberg, W.K., and Dodds, J.B., *Denver Developmental Screening Test,* Denver, Ladoca Project and Publishing Foundation, 1969.

Guilford, J.P., *Psychometric Methods,* New York, McGraw, 1954.

Hambleton, R.K., and Cook, L.L., Latent trait models and their use in the analysis of educational test data, *Journal of Educational Measurement,* 1977, 14, 75-96.

Hieronymous, A.N., and Lindquist, E.F., *Iowa Tests of Basic Skills, Form 5,* Boston, Houghton Mifflin, 1971.

Jastak, J.F., Bijou, S.W., and Jastak, S.R., *Wide Range Achievement Test,* Wilmington, Delaware, Guidance Associates of Delaware, 1965.

Lindquist, E.F., and Feldt, L.S., *Iowa Tests of Educational Development,* Chicago, Science Research Associates, 1972.

Mardell, C., and Goldenberg, D., *Developmental Indicators for the Assessment of Learning,* Highland Park, Illinois, DIAL, 1975.

Rasch, G., *Probalistic Models for Some Intelligence and Attainment Tests,* Copenhagen, Danish Institute for Educational Research, 1960.

Terman, L.M., and Merrill, M.A., *Stanford-Binet Intelligence Scale, Form L-M,* Boston, Houghton Mifflin, 1973.

U.S. Bureau of the Census, Number of inhabitants, *U.S. Census of Population: 1970,* Washington, D.C., U.S. Government Printing Office, 1971.

U.S. Bureau of the Census, Estimates of the population of the United States by age, sex, and race, July 1, 1974 to 1976, *Current Population Reports,* Washington, D.C., U.S. Government Printing Office, 1977.

U.S. Department of Labor, *Employment and Earnings,* January, 1977, 24, 138-139.

Wechsler, D., *Wechsler Adult Intelligence Scale,* New York, Psychological Corporation, 1955.

Wechsler, D., *Wechsler Intelligence Scale for Children—Revised,* New York, Psychological Corporation, 1974.

Woodcock, R.W., *Woodcock Reading Mastery Tests,* Circle Pines, Minnesota, American Guidance Service, 1973.

Woodcock, R.W., and Dahl, M.N., A common scale for the measurement of person ability and test item difficulty, *AGS Paper No. 10,* Circle Pines, Minnesota, American Guidance Service, 1971. (Available from AGS)

Wright, B.D., Sample-free test calibration and person measurement, *Proceedings of the 1967 Invitational Conference on Testing Problems,* Princeton, Educational Testing Service, 1968, 85-101.

Wright, B.D., Solving measurement problems with the Rasch model, *Journal of Educational Measurement,* 1977, 14, 97-116.